

# THE KDV, THE BURGERS, AND THE WHITHAM LIMIT FOR A SPATIALLY PERIODIC BOUSSINESQ MODEL

ROMAN BAUER, WOLF-PATRICK DÜLL, GUIDO SCHNEIDER

IADM, Universität Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart Germany,  
email: forename.lastname@mathematik.uni-stuttgart.de

## CONTENTS

1. Introduction	1
2. The spatially homogeneous case	6
3. Derivation of the amplitude equations	9
4. Estimates for the residual	17
5. The error estimates	19
6. Discussion	24
Appendix A. The inviscid Burgers approximation	25
Appendix B. Higher regularity results	26
Appendix C. Bloch transform on the real line	27
References	28

**ABSTRACT.** We are interested in the Korteweg-de Vries (KdV), the Burgers, and the Whitham limit for a spatially periodic Boussinesq model with non-small contrast. We prove estimates between the KdV, the Burgers, and the Whitham approximation and true solutions of the original system which guarantee that these amplitude equations make correct predictions about the dynamics of the spatially periodic Boussinesq model over the natural time scales of the amplitude equations. The proof is based on Bloch wave analysis and energy estimates. The result is the first justification result of the KdV, the Burgers, and the Whitham approximation for a dispersive PDE posed in a spatially periodic medium of non-small contrast.

## 1. INTRODUCTION

In the long wave limit there exists a zoo of amplitude equations which can be derived via multiple scaling analysis for various dispersive wave systems with conserved quantities. Generically, among these amplitude equations there are only three nonlinear ones which are independent of the small perturbation parameter, namely the Korteweg-de Vries (KdV) equation, the inviscid Burgers equation, and

---

*Date:* January 24, 2017.

the Whitham system. It is the purpose of this paper to discuss the validity of these approximations for a spatially periodic Boussinesq model with non-small contrast.

### 1.1. The formal approximations in the spatially homogenous situation.

The KdV equation occurs as an amplitude equation in the description of small spatially and temporally modulations of long waves in various dispersive wave systems. Examples are the water wave problem or equations from plasma physics, cf. [3]. For the Boussinesq equation

$$(1) \quad \partial_t^2 u(x, t) = \partial_x^2 u(x, t) - \partial_x^4 u(x, t) + \partial_x^2 (u(x, t)^2),$$

with  $x \in \mathbb{R}$ ,  $t \in \mathbb{R}$ , and  $u(x, t) \in \mathbb{R}$ , by the ansatz

$$(2) \quad u(x, t) = \varepsilon^2 A(X, T),$$

where  $X = \varepsilon(x - t)$ ,  $T = \varepsilon^3 t$ ,  $A(X, T) \in \mathbb{R}$ , and  $0 < \varepsilon \ll 1$  a small perturbation parameter, the KdV equation

$$(3) \quad \partial_T A = \frac{1}{2} \partial_X^3 A - \frac{1}{2} \partial_X (A^2)$$

can be derived by inserting (2) into (1) and by equating the coefficients in front of  $\varepsilon^6$  to zero. This ansatz can be generalized to

$$(4) \quad u(x, t) = \varepsilon^\alpha A(X, T),$$

where  $X = \varepsilon(x - t)$ ,  $T = \varepsilon^{1+\alpha} t$ , and  $A(X, T) \in \mathbb{R}$ , with  $\alpha > 0$ . For  $\alpha > 2$  the Airy equation  $\partial_T A = \frac{1}{2} \partial_X^3 A$  occurs. The KdV equation is recovered for  $\alpha = 2$ , and for  $\alpha \in (0, 2)$  the inviscid Burgers equation

$$(5) \quad \partial_T A = -\frac{1}{2} \partial_X (A^2)$$

is obtained. There is another long wave limit which leads to an  $\varepsilon$ -independent non-trivial amplitude equation. With the ansatz

$$(6) \quad u(x, t) = U(X, T),$$

where  $X = \varepsilon x$ ,  $T = \varepsilon t$ , and  $U(X, T) \in \mathbb{R}$ , we obtain

$$(7) \quad \partial_T^2 U = \partial_X^2 U + \partial_X (U^2)$$

which can be written as a system of conservation laws

$$(8) \quad \partial_T U = \partial_X V, \quad \partial_T V = \partial_X U + \partial_X (U^2).$$

In the following both, (7) and (8), are called the Whitham system, cf. [26].

**1.2. Justification by error estimates.** Estimates that the formal KdV approximation and true solutions of the original system stay close together over the natural KdV time scale are a non-trivial task since solutions of order  $\mathcal{O}(\varepsilon^2)$  have to be shown to be existent on an  $\mathcal{O}(1/\varepsilon^3)$  time scale. For (1) an approximation result is formulated as follows.

**Theorem 1.1.** *Let  $s \geq 0$  and let  $A \in C([0, T_0], H^{5+s})$  be a solution of the KdV equation (3). Then there exist  $\varepsilon_0 > 0$  and  $C > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have solutions  $u$  of (1) with*

$$\sup_{t \in [0, T_0/\varepsilon^3]} \|u(\cdot, t) - \varepsilon^2 A(\varepsilon(\cdot - t), \varepsilon^3 t)\|_{H^{1+s}} \leq C \varepsilon^{7/2}.$$

There are two fundamentally different approaches to prove such an approximation result. For analytic initial conditions of the KdV equation a Cauchy-Kowalevskaya based approach can be chosen, see [19] with the comments given in [22] for the water wave problem. Working in spaces of analytic functions gives some artificial smoothing which allows to gain the missing order w.r.t.  $\varepsilon$  between the inverse of the amplitude of  $\mathcal{O}(\varepsilon^2)$  and the time scale of  $\mathcal{O}(1/\varepsilon^3)$  via the derivative in front of the nonlinear terms in the KdV equation. This approach is very robust and works without a detailed analysis of the underlying problem, cf. [5] for another example, but gives not optimal results.

For initial conditions in Sobolev spaces the underlying idea to gain such estimates is conceptually rather simple, namely the construction of a suitable chosen energy which include the terms of order  $\mathcal{O}(\varepsilon^2)$  in the equation for the error, such that for the energy finally  $\mathcal{O}(\varepsilon^3 t)$  growth rates occur. However, the method is less robust since for every single original system a different energy occurs and the major difficulty is the construction of this energy. Estimates that the formal KdV approximation and true solutions of the different formulations of the water wave problem stay close together over the natural KdV time scale have been shown for instance in [10, 23, 24, 1, 14] using this approach. Another example is the justification of the KdV approximation for modulations of periodic waves in the NLS equation, cf. [6]. For (1) the energy approach is rather short and very instructive for the subsequent analysis. Therefore, we recall it in Section 2.

Interestingly, it turns out that the proofs given for the KdV approximations transfer more or less line for line into proofs for the justification of the inviscid Burgers equation and of the Whitham system. Since only the scaling has to be adapted, whenever a KdV approximation result holds also an inviscid Burgers and Whitham approximation result can be established. This will be explained in detail in Section 2.

As above such approximation results are a non-trivial task since solutions of order  $\mathcal{O}(\varepsilon^\alpha)$  have to be shown to be existent on an  $\mathcal{O}(\varepsilon^{1+\alpha})$  time scale. For the inviscid Burgers equation the formulation of the approximation result goes along the lines of Theorem 1.1. However, due to the notational complexity in achieving

in general the estimates for the residual (the terms which do not cancel after inserting the approximation into (9)), in Remark 2.3 we restrict ourselves to the case  $\alpha = 1$ .

**Theorem 1.2.** *Let  $s \geq 0$ ,  $\alpha = 1$  and let  $A \in C([0, T_0], H^{3+s})$  be a solution of the inviscid Burgers equation (5). Then there exist  $\varepsilon_0 > 0$  and  $C > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have solutions  $u$  of (1) with*

$$\sup_{t \in [0, T_0/\varepsilon^2]} \|u(\cdot, t) - \varepsilon A(\varepsilon(\cdot - t), \varepsilon^2 t)\|_{H^{1+s}} \leq C\varepsilon^{(3+2\alpha)/2}.$$

Since for the Whitham approximation solutions of order  $\mathcal{O}(1)$  are considered some smallness condition is needed such that the used energy allows us to estimate the associated Sobolev norm.

For (1) a possible Whitham approximation result is formulated as follows.

**Theorem 1.3.** *Let  $s \geq 0$ . There exists a  $C_1 > 0$  such that the following holds. Let  $U \in C([0, T_0], H^{3+s})$  be a solution of (7) with*

$$\sup_{T \in [0, T_0]} \|U(\cdot, T)\|_{H^{3+s}} \leq C_1.$$

*Then there exist  $\varepsilon_0 > 0$  and  $C > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have solutions  $u$  of (1) with*

$$\sup_{t \in [0, T_0/\varepsilon]} \|u(\cdot, t) - U(\varepsilon \cdot, \varepsilon t)\|_{H^{1+s}} \leq C\varepsilon^{3/2}.$$

The Whitham system for the water wave problem coincides with the shallow water wave equations which have been justified for the water wave problem without surface tension in [20, 17]. A Whitham approximation result that the periodic wave trains of the NLS equation are approximated by the Whitham system can be found in [13].

**1.3. The spatially periodic situation.** The last years have seen some first attempts to justify the KdV equation in periodic media. It has been justified in [17] for the water wave problem over a periodic bottom in the KdV scaling, i.e., with long wave oscillations of the bottom of magnitude  $\mathcal{O}(\varepsilon^2)$  varying on a spatial scale of order  $\mathcal{O}(\varepsilon^{-1})$ . The same result can be found in [9] where general bottom topographies of small amplitude have been handled. The result is based on [8] where other amplitude systems have been justified. This situation can be handled as perturbation of the spatially homogeneous case.

In case of oscillations of the bottom of magnitude  $\mathcal{O}(1)$  varying on a spatial scale of order  $\mathcal{O}(1)$ , no approximation result can be found in the existing literature. As a first attempt to solve this question for the water wave problem we consider a spatially periodic Boussinesq equation

$$(9) \quad \begin{aligned} \partial_t^2 u(x, t) = & \partial_x(a(x)\partial_x u(x, t)) \\ & - \partial_x^2(b(x)\partial_x^2 u(x, t)) + \partial_x(c(x)\partial_x(u(x, t)^2)), \end{aligned}$$

with  $x \in \mathbb{R}$ ,  $t \geq 0$ ,  $u(x, t) \in \mathbb{R}$ , and smooth  $x$ -dependent  $2\pi$ -spatially periodic coefficients  $a$ ,  $b$ , and  $c$  satisfying

$$\inf_{x \in \mathbb{R}} a(x) > 0 \quad \text{and} \quad \inf_{x \in \mathbb{R}} b(x) > 0.$$

For this equation we derive the KdV equation by making a Bloch mode expansion of (9). The KdV approximation describes the modes which in Figure 1 are contained in the circles. We prove an approximation result which is formulated in Theorem 5.1. It guarantees that the KdV equation makes correct predictions about the dynamics of the spatially periodic Boussinesq model (9) over the natural KdV time scale. The presented result is the first justification result of the KdV approximation for a dispersive nonlinear PDE posed in a spatially periodic medium of non-small contrast. For linear systems this limit has been considered independently in [11, 12].

In order to make the residual small an improved approximation has to be constructed. Since this construction is not the main purpose of this paper we additionally assume

**(SYM)** the coefficient functions

$$a = a(x), \quad b = b(x), \quad \text{and} \quad c = c(x) \quad \text{are even w.r.t. } x.$$

As in the spatially homogeneous situation it turns out that the proof given for the KdV approximation transfers more or less line for line into proofs for the justification of the approximation via the inviscid Burgers equation and of the Whitham system. The associated approximation results are formulated in Theorem 5.2 and Theorem 5.3.

The paper was originally intended as the next step in generalizing a method which has been developed in [7] for the justification of the KdV approximation in situations when the KdV modes are resonant to other long wave modes. The method had already successfully been applied in justifying the KdV approximation for the poly-atomic FPU problem in [4]. The qualitative difference in justifying the KdV equation for the spatially periodic Boussinesq equation in contrast to [7, 4] is that for fixed Bloch respectively Fourier wave number the presented problem is infinite dimensional. [7, 4] corresponds to the middle panel of Figure 1 where the spatially periodic Boussinesq equation corresponds to the right panel of Figure 1. As a consequence the normal form transform which is a major part of the proofs of [7, 4] would be more demanding from an analytic point of view. In the justification of the Whitham system with the approach of [7, 4] infinitely many normal form transforms have to be performed [15].

Interestingly, for the spatially periodic Boussinesq equation (9) there exists an energy in physical space which allowed us to incorporate the normal form transforms into the energy estimates. This energy approach is presented in the following.

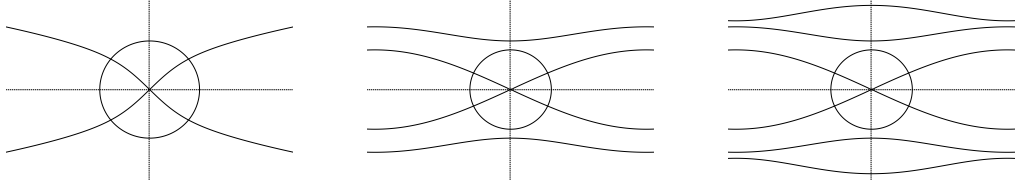


FIGURE 1. The left panel shows the curves of eigenvalues over the Fourier wave numbers as it appears for the water wave problem [10, 23, 24, 1, 17, 14]. The middle panel shows the finitely many curves of eigenvalues as they appear for instance for the poly-atomic FPU system [7, 4]. The right panel shows the infinitely many curves of eigenvalues over the Bloch wave numbers as it appears for the spatially periodic Boussinesq model (9), the water wave problem over a periodic bottom topography, or for the linearization around a periodic wave in dispersive systems. Since the Fourier transform of  $\varepsilon^2 A(\varepsilon x)$  is given by  $\varepsilon^2 \varepsilon^{-1} \hat{A}(x/\varepsilon)$  the KdV equations describe the modes at the wave numbers  $k = 0$  with the vanishing eigenvalues which are contained in the circles. One of the two curves in the circle describes wave packets moving to the left, the other curve wave packets moving to the right.

**Notation.** Constants which can be chosen independently of the small perturbation parameter  $0 < \varepsilon \ll 1$  are denoted with the same symbol  $C$ . We write  $\int$  for  $\int_{-\infty}^{\infty}$ . The Fourier transform of a function  $u$  is denoted with  $\hat{u}$ . The Bloch transform of a function  $u$  is denoted with  $\tilde{u}$  and this tool is recalled in Appendix C. We introduce the norm  $\|\cdot\|_{L_s^2}$  by

$$\|\hat{u}\|_{L_s^2}^2 = \int |\hat{u}(k)|^2 (1 + k^2)^s dk$$

and define the Sobolev norm  $\|u\|_{H^s} = \|\hat{u}\|_{L_s^2}$ , but use also equivalent versions.

**Acknowledgement.** The authors are grateful to Florent Chazel for helping us to understand the existing literature. Moreover, we would like to thank Martina Chirilus-Bruckner for a number of helpful discussions. The paper is partially supported by the Deutsche Forschungsgemeinschaft (DFG) under the grant Schn520/9-1.

## 2. THE SPATIALLY HOMOGENEOUS CASE

It is the goal of this section to give a simple proof for Theorem 1.1, Theorem 1.2, and Theorem 1.3 using the energy method. The proof will be the basis of the subsequent analysis. All three cases can be handled with the same approach.

The residual

$$\text{Res}(u) = -\partial_t^2 u(x, t) + \partial_x^2 u(x, t) - \partial_x^4 u(x, t) + \partial_x^2 (u(x, t)^2)$$

quantifies how much a function  $u$  fails to satisfy the Boussinesq model (1). For the KdV approximation (2) abbreviated with  $\varepsilon^2\Psi$  we find

$$\begin{aligned}\text{Res}(\varepsilon^2\Psi) &= -\varepsilon^4 c^2 \partial_X^2 A - 2\varepsilon^6 \partial_T \partial_X A - \varepsilon^8 \partial_T^2 A \\ &\quad + \varepsilon^4 \partial_X^2 A - \varepsilon^6 \partial_X^4 A + \varepsilon^6 \partial_X^2 (A^2) \\ &= -\varepsilon^8 \partial_T^2 A\end{aligned}$$

if we choose  $A$  to satisfy the KdV equation (3). Therefore, we have

**Lemma 2.1.** *Let  $s \geq 0$  and let  $A \in C([0, T_0], H^{5+s})$  be a solution of the KdV equation (3). Then there exist  $\varepsilon_0 > 0$ ,  $C_{\text{res}}$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have*

$$\sup_{t \in [0, T_0/\varepsilon^3]} \|\partial_x^{s-1} \text{Res}(\varepsilon^2\Psi(\cdot, t, \varepsilon))\|_{L^2} \leq C_{\text{res}} \varepsilon^{(13+2s)/2}.$$

*Proof.* Using the KdV equation allows us to write

$$\begin{aligned}4\partial_T^2 A &= -2\partial_T(\partial_X^3 A + \partial_X(A^2)) = -2(\partial_X^3 \partial_T A + 2\partial_X(A\partial_T A)) \\ &= \partial_X^3(\partial_X^3 A + \partial_X(A^2)) + 2\partial_X(A(\partial_X^3 A + \partial_X(A^2))).\end{aligned}$$

This shows that  $A(\cdot, T) \in H^6$  is necessary to estimate the residual in  $L^2$ . The formal error of order  $\mathcal{O}(\varepsilon^8)$  is reduced by a factor  $\varepsilon^{-1/2}$  due to the scaling properties of the  $L^2$ -norm. Moreover, due to the representation of  $\partial_T^2 A$  as a spatial derivative, below, we can apply  $\partial_x^{-1} = \varepsilon^{-1} \partial_X^{-1}$  to the residual terms which however loses another factor  $\varepsilon^{-1}$ .  $\square$

Similarly, for the Whitham approximation (6) abbreviated with  $\varepsilon^2\Psi$  we find  $\text{Res}(\Psi) = -\varepsilon^4 \partial_X^4 U$  if we choose  $U$  to satisfy the Whitham system (7). Hence, for an estimate in  $L^2$  we need  $U \in H^4$ . Exactly as above we have

**Lemma 2.2.** *Let  $s \geq 0$  and let  $A \in C([0, T_0], H^{3+s})$  be a solution of the Whitham system (7). Then there exist  $\varepsilon_0 > 0$ ,  $C_{\text{res}}$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have*

$$\sup_{t \in [0, T_0/\varepsilon]} \|\partial_x^{s-1} \text{Res}(\Psi(\cdot, t, \varepsilon))\|_{L^2} \leq C_{\text{res}} \varepsilon^{(5+2s)/2}.$$

**Remark 2.3.** For the inviscid Burgers equation the residual becomes too large with the simple ansatz (2). However, by adding higher order terms to the approximation (2), with a slight abuse of notation this approximation is again called  $\varepsilon^\alpha\Psi$ , one can always achieve

$$\sup_{t \in [0, T_0/\varepsilon^{1+\alpha}]} \|\text{Res}(\varepsilon^\alpha\Psi(\cdot, t, \varepsilon))\|_{L^2} \leq C_{\text{res}} \varepsilon^{(7+4\alpha)/2}$$

and

$$\sup_{t \in [0, T_0/\varepsilon^{1+\alpha}]} \|\partial_x^{-1} \text{Res}(\varepsilon^\alpha\Psi(\cdot, t, \varepsilon))\|_{L^2} \leq C_{\text{res}} \varepsilon^{(5+4\alpha)/2}.$$

See Appendix A where we prove these estimates for  $\alpha = 1$  and explain that the number of additional terms goes to infinity for  $\alpha \rightarrow 0$  and  $\alpha \rightarrow 2$ .

From this point on the remaining estimates can be handled exactly the same. The case  $\alpha = 0$  corresponds to the Whitham approximation and the case  $\alpha = 2$  to the KdV approximation. The difference  $\varepsilon^{(3+2\alpha)/2}R = u - \varepsilon^\alpha \Psi$  satisfies

$$(10) \quad \partial_t^2 R = \partial_x^2 R - \partial_x^4 R + 2\varepsilon^\alpha \partial_x^2(\Psi R) + \varepsilon^{(3+2\alpha)/2} \partial_x^2(R^2) + \varepsilon^{-(3+2\alpha)/2} \text{Res}(\varepsilon^2 \Psi).$$

We multiply the error equation (10) with  $-\partial_t \partial_x^{-2} R$  which is defined via its Fourier transform w.r.t.  $x$ , namely via  $\widehat{\partial_x^{-2} R}(k) = \frac{1}{ik} \widehat{R}(k)$ , integrate it w.r.t.  $x$ , and find

$$\begin{aligned} - \int (\partial_t \partial_x^{-2} R) \partial_t^2 R dx &= \partial_t \int (\partial_t \partial_x^{-1} R)^2 dx / 2, \\ - \int (\partial_t \partial_x^{-2} R) \partial_x^2 R dx &= -\partial_t \int R^2 dx / 2, \\ \int (\partial_t \partial_x^{-2} R) \partial_x^4 R dx &= -\partial_t \int (\partial_x R)^2 dx / 2, \\ - \int (\partial_t \partial_x^{-2} R) \partial_x^2(\Psi R) dx &= - \int (\partial_t R) \Psi R dx \\ &= -\partial_t \int \Psi R^2 dx / 2 + \varepsilon \int (\partial_\tau \Psi) R^2 dx, \\ - \int (\partial_t \partial_x^{-2} R) \partial_x^2(R^2) dx &= - \int (\partial_t R) R^2 dx = -\frac{1}{3} \partial_t \int R^3 dx, \\ - \int (\partial_t \partial_x^{-2} R) \text{Res}(\varepsilon^2 \Psi) dx &= \int (\partial_t \partial_x^{-1} R) \partial_x^{-1} \text{Res}(\varepsilon^2 \Psi) dx. \end{aligned}$$

We can estimate

$$\begin{aligned} \left| \int (\partial_t \partial_x^{-1} R) \partial_x^{-1} \text{Res}(\varepsilon^2 \Psi) dx \right| &\leq \| \partial_t \partial_x^{-1} R \|_{L^2} \| \partial_x^{-1} \text{Res}(\varepsilon^2 \Psi) \|_{L^2}, \\ \left| \int (\partial_\tau \Psi) R^2 dx \right| &\leq \| \partial_\tau \Psi \|_{L^\infty} \| R \|_{L^2}^2. \end{aligned}$$

For the energy

$$E = \int (\partial_t \partial_x^{-1} R)^2 + R^2 + (\partial_x R)^2 + 2\varepsilon^\alpha \Psi R^2 + 2\varepsilon^{(3+2\alpha)/2} R^3 / 3 dx$$

the following holds. In case  $\alpha > 0$  we have that for all  $M > 0$  there exist  $C_1, \varepsilon_1 > 0$  such that for all  $\varepsilon \in (0, \varepsilon_1)$  we have

$$\| R \|_{H^1} \leq C_1 E^{1/2}$$

as long as  $E \leq M$ . In case  $\alpha = 0$  the energy  $E$  is an upper bound for the squared  $H^1$ -norm for  $\| \Psi \|_{L^\infty}$  sufficiently small, but independent of  $0 < \varepsilon \ll 1$ . Therefore,  $E$  satisfies the inequality

$$\begin{aligned} (11) \quad \frac{dE}{dt} &\leq C\varepsilon^{1+\alpha} E + C\varepsilon^{(3+2\alpha)/2} E^{3/2} + C\varepsilon^{1+\alpha} E^{1/2} \\ &\leq 2C\varepsilon^{1+\alpha} E + C\varepsilon^{(3+2\alpha)/2} E^{3/2} + C\varepsilon^{1+\alpha}, \end{aligned}$$



with a constant  $C$  independent of  $\varepsilon \in (0, \varepsilon_1)$ . Under the assumption that  $C\varepsilon^{1/2}E^{1/2} \leq 1$  we obtain

$$\frac{dE}{dt} \leq (2C + 1)\varepsilon^{1+\alpha}E + C\varepsilon^{1+\alpha}.$$

Gronwall's inequality immediately gives the bound

$$\sup_{t \in [0, T_0/\varepsilon^{1+\alpha}]} E(t) = CT_0 e^{(2C+1)T_0} =: M = \mathcal{O}(1).$$

Finally choosing  $\varepsilon_2 > 0$  so small that  $C\varepsilon_2^{1/2}M^{1/2} \leq 1$  gives the required estimate for all  $\varepsilon \in (0, \varepsilon_0)$  with  $\varepsilon_0 = \min(\varepsilon_1, \varepsilon_2) > 0$  in all three cases.

**Remark 2.4.** The Boussinesq model (1) is a semilinear dispersive system and so there is the local existence and uniqueness of solutions. The variation of constant formula associated to the first order system for the variables  $u$  and  $\partial_t(\partial_x^4 - \partial_x^2)^{-1/2}u$  is a contraction in the space  $C([-T_*, T_*], H^\theta \times H^\theta)$  for every  $\theta > 1/2$  if  $T_* > 0$  is sufficiently small. The local existence and uniqueness of solutions combined with the previous estimates for instance yields the existence and uniqueness of solutions for all  $t \in [0, T_0/\varepsilon^3]$  in the KdV case and all  $t \in [0, T_0/\varepsilon]$  in the Whitham case.

### 3. DERIVATION OF THE AMPLITUDE EQUATIONS

In this section we come back to the spatially periodic situation. The derivation of the amplitude equations is less obvious than in the spatially homogeneous case. In order to derive the amplitude equations we expand (9) into the eigenfunctions of the linear problem. As in [2] after this expansion we are back in the spatially homogeneous set-up except that Fourier transform has been replaced by the Bloch transform.

**3.1. Spectral properties.** The linearized problem

$$(12) \quad \partial_t^2 u(x, t) = \partial_x(a(x)\partial_x u(x, t)) - \partial_x^2(b(x)\partial_x^2 u(x, t))$$

is solved by so called Bloch modes

$$u(x, t) = w(x)e^{ilx}e^{i\omega t},$$

with  $w$  being  $2\pi$ -periodic w.r.t.  $x$  satisfying

$$-(\partial_x + il)(a(x)(\partial_x + il)w(x)) + (\partial_x + il)^2(b(x)(\partial_x + il)^2w(x)) = \omega^2 w(x).$$

The left hand side defines a self-adjoint elliptic operator  $L_l(\partial_x) : H^{\theta+4} \rightarrow H^\theta$ . Hence, for fixed  $l$  there exists a countable set of eigenvalues  $\lambda_n(l)$ , with  $n \in \mathbb{N}$ , ordered such that  $\lambda_{n+1}(l) \geq \lambda_n(l)$ , with associated eigenfunctions  $w_n(x, l)$ .

**Lemma 3.1.** *For  $l = 0$  the operator  $L_0(\partial_x)$  possesses the simple eigenvalue  $\lambda_1(0) = 0$  associated to the eigenfunction  $\tilde{w}_1(0, x) = 1$ .*

**Proof.** Obviously we have  $L_0(\partial_x)1 = 0$ . Moreover, we have

$$(w, L_0(\partial_x)w)_{L^2} = \int_{-1/2}^{1/2} a(x)(\partial_x w(x))^2 dx + \int_{-1/2}^{1/2} b(x)(\partial_x^2 w(x))^2 dx \geq 0.$$

Hence  $L_0(\partial_x)w = 0$  implies  $\partial_x w = 0$ . From the  $2\pi$ -periodicity it follows  $w = \text{const}$ . Hence  $\lambda_1(0) = 0$  is a simple eigenvalue.  $\square$

It is well known that the curves  $l \mapsto \lambda_n(l)$  and  $l \mapsto \tilde{w}_n(l, \cdot)$  are smooth w.r.t.  $l$  for simple eigenvalues. Hence, there exists a  $\delta_0 > 0$  such that for  $l \in [-\delta_0, \delta_0]$  the smallest eigenvalue  $\lambda_1(l)$  is separated from the rest of the spectrum. Since  $L_l(\partial_x)$  is self-adjoint and positive-definite for all  $l$  we have  $\lambda_1(l) \geq 0$  for all  $l$ . In the KdV equation only odd and in the Whitham system only even spatial derivatives occur. This is a consequence of the following lemma.

**Lemma 3.2.** *The curve  $l \mapsto \lambda_1(l)$  for  $l \in [-\delta_0, \delta_0]$  is an even real-valued function. The associated eigenfunctions satisfy  $\tilde{w}_1(l, x) = \overline{\tilde{w}_1(-l, x)}$ . Under the assumption that the coefficient functions  $a$  and  $b$  are even, the eigenfunctions possess an expansion*

$$\tilde{w}_1(l, x) = \sum_{j=0}^{\infty} (il)^j g_j(x),$$

with  $g_0(x) = 1$ ,  $\int_0^{2\pi} g_j(x) dx = 0$  for  $j \geq 1$ ,

$$g_{2j}(x) = g_{2j}(-x) \in \mathbb{R} \quad \text{and} \quad g_{2j+1}(x) = -g_{2j+1}(-x) \in \mathbb{R}.$$

**Proof.** The first two statements follow from the fact that for fixed  $l$  the operator  $L_l(\partial_x)$  is self-adjoint and from the fact that (9) is a real problem. For  $(il)^0$  we obtain

$$-\partial_x(a(x)\partial_x g_0(x)) + \partial_x^2(b(x)\partial_x^2 g_0(x)) = 0$$

which is, as we already know, uniquely been solved by  $g_0(x) = 1$ . For  $(il)^1$  we obtain

$$-\partial_x(a(x)\partial_x g_1(x)) + \partial_x^2(b(x)\partial_x^2 g_1(x)) - \partial_x a(x) = 0.$$

The term  $\partial_x a(x)$  is odd. The subspace of odd functions is invariant for the differential operator  $L_0(\partial_x) = -\partial_x(a(x)(\partial_x \cdot)) + \partial_x^2(b(x)\partial_x^2 \cdot)$ . Moreover in this subspace its spectrum is bounded away from zero such that this equation possesses a unique odd solution  $g_1 = g_1(x)$ . For  $(il)^2$  we obtain

$$-\partial_x(a(x)\partial_x g_2(x)) + \partial_x^2(b(x)\partial_x^2 g_2(x)) + 1 + f_2(x) = 1$$

with  $f_2(x)$  an even function depending on  $a$ ,  $b$ ,  $g_0$ , and  $g_1$  and possessing vanishing mean value. In the subspace of vanishing mean value the differential operator  $L_0(\partial_x)$  possesses spectrum which is bounded away from zero such that this equation possesses a unique even solution  $g_2 = g_2(x)$ . With the same arguments the next orders with the stated properties can be computed. The convergence of the series in a neighborhood of  $l = 0$  in  $H^\theta$  for every  $\theta \geq 0$  follows from the smoothness of

the curve of simple eigenfunctions w.r.t.  $l$  and the smoothness of the coefficient functions  $a$ ,  $b$ , and  $c$  w.r.t.  $x$ .  $\square$

The KdV equation, the inviscid Burgers equation, and the Whitham system describe the modes associated to the curve  $\lambda_1$  close to  $l = 0$ . Therefore, in order to derive these amplitude equations we consider the Bloch transform

$$u(x, t) = \int_{-1/2}^{1/2} \tilde{u}(l, x, t) e^{ilx} dx$$

of (9), namely

$$(13) \quad \partial_t^2 \tilde{u}(l, x, t) = -L_l(\partial_x) \tilde{u}(l, x, t) + N_l(\partial_x)(\tilde{u})(l, x, t)$$

where

$$N_l(\partial_x)(\tilde{u})(l, x, t) = (\partial_x + il)(c(x)(\partial_x + il) \int_{-1/2}^{1/2} \tilde{u}(l - m, x, t) \tilde{u}(m, x, t) dm.$$

Then we make the ansatz

$$\tilde{u}(l, x, t) = \chi_{[-\delta_0/2, \delta_0/2]}(l) \tilde{u}_1(l, t) \tilde{w}_1(l, x) + \tilde{v}(l, x, t)$$

with

$$\int_0^{2\pi} \overline{\tilde{w}_1(l, x)} \tilde{v}(l, x, t) dx = 0$$

for  $l \in [-\delta_0/2, \delta_0/2]$  and find

$$\begin{aligned} \partial_t^2 \tilde{u}_1(l, t) &= -\lambda_1(l) \tilde{u}_1(l, t) + P_c(l) N_l(\partial_x)(\tilde{u})(l, t), \\ \partial_t^2 \tilde{v}(l, x, t) &= -L_l(\partial_x) \tilde{v}(l, x, t) + P_s(l) N_l(\partial_x)(\tilde{u})(l, x, t), \end{aligned}$$

where

$$\begin{aligned} (P_c \tilde{u})(l, t) &= \frac{1}{2\pi} \chi_{[-\delta_0/2, \delta_0/2]}(l) \int_0^{2\pi} \overline{\tilde{w}_1(l, x)} \tilde{u}(l, x, t) dx, \\ (P_s \tilde{u})(l, x, t) &= \tilde{u}(l, x, t) - (P_c \tilde{u})(l, t) \tilde{w}_1(l, x). \end{aligned}$$

All amplitude equations which we have in mind can be derived in a very similar way. They describe the evolution of the  $\tilde{u}_1$  modes which are concentrated in an  $\mathcal{O}(\varepsilon)$  neighborhood of the Bloch wave number  $l = 0$ . In all three cases we make an ansatz

$$(14) \quad \tilde{u}_1(l, t) = \varepsilon^{-1} \varepsilon^\alpha \chi_{[-\delta_0/4, \delta_0/4]}(\frac{l}{\varepsilon}) \hat{A}(\frac{l}{\varepsilon}, \varepsilon^{1+\alpha} t) e^{ilct}$$

with  $\alpha = 2$  and  $c > 0$  for the KdV approximation,  $\alpha \in (0, 2)$  and  $c > 0$  for the inviscid Burgers approximation, and  $\alpha = 0$  and  $c = 0$  for the Whitham approximation, cf. the text below Figure 1. The amplitude  $\hat{A}$  will be defined in Fourier

space and the cut-off function  $\chi_{[-\delta_0/4, \delta_0/4]}(\frac{l}{\varepsilon})$  allows to transfer  $\widehat{A}$  into Bloch space. In the following we use the abbreviation

$$(15) \quad \widetilde{A}(\frac{l}{\varepsilon}, \varepsilon^{1+\alpha}t) = \chi_{[-\delta_0/4, \delta_0/4]}(\frac{l}{\varepsilon}) \widehat{A}(\frac{l}{\varepsilon}, \varepsilon^{1+\alpha}t).$$

For each of the three approximations we have to derive the associated amplitude equation and to compute and estimate the residual terms

$$\text{Res}(\widetilde{u})(l, x, t) = -\partial_t^2 \widetilde{u}(l, x, t) - L_l(\partial_x) \widetilde{u}(l, x, t) + N_l(\partial_x)(\widetilde{u})(l, x, t).$$

**3.2. Derivation of the KdV and the inviscid Burgers equation.** The amplitude equations which we have in mind have derivatives in front of the nonlinear terms. Hence before deriving these equations we need to prove a number of properties about the nonlinear terms. We introduce kernels  $s_{11}^1(l, l-m, m), \dots, s_{vv}^v(l, l-m, m)$  by

$$\begin{aligned} (P_c N_l(\partial_x)(\widetilde{u}))(l, t) &= \int_{-1/2}^{1/2} s_{11}^1(l, l-m, m) \widetilde{u}_1(l-m, t) \widetilde{u}_1(m, t) dm \\ &+ \int_{-1/2}^{1/2} s_{1v}^1(l, l-m, m) \widetilde{u}_1(l-m, t) \widetilde{v}(m, x, t) dm \\ &+ \int_{-1/2}^{1/2} s_{v1}^1(l, l-m, m) \widetilde{v}(l-m, x, t) \widetilde{u}_1(m, t) dm \\ &+ \int_{-1/2}^{1/2} s_{vv}^1(l, l-m, m) \widetilde{v}(l-m, x, t) \widetilde{v}(m, x, t) dm \end{aligned}$$

and

$$\begin{aligned} (P_s N_l(\partial_x)(\widetilde{u}))(l, x, t) &= \int_{-1/2}^{1/2} s_{11}^v(l, l-m, m) \widetilde{u}_1(l-m, t) \widetilde{u}_1(m, t) dm \\ &+ \int_{-1/2}^{1/2} s_{1v}^v(l, l-m, m) \widetilde{u}_1(l-m, t) \widetilde{v}(m, x, t) dm \\ &+ \int_{-1/2}^{1/2} s_{v1}^v(l, l-m, m) \widetilde{v}(l-m, x, t) \widetilde{u}_1(m, t) dm \\ &+ \int_{-1/2}^{1/2} s_{vv}^v(l, l-m, m) \widetilde{v}(l-m, x, t) \widetilde{v}(m, x, t) dm. \end{aligned}$$

For the derivation of the KdV and the Burgers equation we need

**Lemma 3.3.** *We have*

$$|s_{11}^1(l, l-m, m) - \nu_2 l^2| \leq C|l|(l^2 + (l-m)^2 + m^2),$$

where

$$(16) \quad \nu_2 = -\frac{1}{2\pi} \int_0^{2\pi} c(x)(1 + \partial_x g_1(x))^2 dx.$$

**Proof.** Due to Lemma 3.2 we have

$$(17) \quad \tilde{w}_1(l, x) = 1 + ilg_1(x) + \mathcal{O}(l^2)$$

where  $g_1(x) \in \mathbb{R}$  with  $\int_0^{2\pi} g_1(x) dx = 0$ . This expansion yields

$$\begin{aligned} & 2\pi s_{11}^1(l, l-m, m) \\ = & \int_0^{2\pi} \overline{\tilde{w}_1(l, x)} (\partial_x + il)(c(x)(\partial_x + il)(\tilde{w}_1(l-m, x)\tilde{w}_1(m, x))) dx \\ = & \int_0^{2\pi} (1 - ilg_1(x) + \mathcal{O}(l^2))(\partial_x + il)(c(x)(\partial_x + il) \\ & \quad \times ((1 + i(l-m)g_1(x) + \mathcal{O}((l-m)^2))(1 + img_1(x) + \mathcal{O}(m^2)))) dx \\ = & - \int_0^{2\pi} c(x)((\partial_x - il)(1 - ilg_1(x) + \mathcal{O}(l^2))((\partial_x + il) \\ & \quad \times ((1 + i(l-m)g_1(x) + \mathcal{O}((l-m)^2))(1 + img_1(x) + \mathcal{O}(m^2)))) dx \\ = & - \int_0^{2\pi} c(x)(-il - il\partial_x g_1(x) + \mathcal{O}(l^2)) \\ & \quad \times ((\partial_x + il)((1 + ilg_1(x) + \mathcal{O}((l-m)^2 + m^2))) dx \\ = & - \int_0^{2\pi} c(x)(-il - il\partial_x g_1(x) + \mathcal{O}(l^2)) \\ & \quad \times (il + il\partial_x g_1(x) + \mathcal{O}((l-m)^2 + m^2)) dx \\ = & \nu_2 l^2 + \mathcal{O}(|l|(l^2 + (l-m)^2 + m^2)). \end{aligned}$$

We remark already at this point that due to the fact that  $a$ ,  $b$ , and  $c$  are assumed to be even we have for symmetry reasons that the higher order terms are not only  $\mathcal{O}(|l|(l^2 + (l-m)^2 + m^2))$ , but  $\mathcal{O}(l^4 + (l-m)^4 + m^4)$ . See below.  $\square$

The following derivation of amplitude equations in Fourier or Bloch space is straightforward and documented in various papers. We refer to [25, Chapter 5] for an introduction.

**3.2.1. The KdV equation.** We start with the KdV approximation  $\varepsilon^2 \Psi$  which is defined via (14) for  $\alpha = 2$  and which is inserted into  $\text{Res}(\tilde{u})$ . We find with  $\tilde{u}_1(l, t) =$

$\varepsilon \tilde{A}(K, T) \mathbf{E}$ ,  $\mathbf{E} = e^{i\varepsilon Kct}$ ,  $T = \varepsilon^3 t$ , and  $l = \varepsilon K$  that

$$\begin{aligned}
P_c(\text{Res}(\tilde{u}))(l, t) &= -\partial_t^2 \tilde{u}_1(l, t) - \lambda_1(l) \tilde{u}_1(l, t) \\
&\quad + \int_{-1/2}^{1/2} s_{11}^1(l, l-m, m) \tilde{u}_1(l-m, t) \tilde{u}_1(m, t) dm \\
&= \varepsilon^3 c^2 K^2 \tilde{A}(K, T) \mathbf{E} - 2\varepsilon^5 icK (\partial_T \tilde{A}(K, T)) \mathbf{E} - \varepsilon^7 (\partial_T^2 \tilde{A}(K, T)) \mathbf{E} \\
&\quad - \varepsilon^3 \lambda_1''(0) K^2 \tilde{A}(K, T) \mathbf{E} / 2 - \varepsilon^5 \lambda_1''''(0) K^4 \tilde{A}(K, T) \mathbf{E} / 24 + \mathcal{O}(\varepsilon^7) \\
&\quad + \varepsilon^5 \int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \nu_2 K^2 \tilde{A}(K-M, T) \tilde{A}(M, T) dM \mathbf{E} + \mathcal{O}(\varepsilon^6).
\end{aligned}$$

If  $\hat{A}(\cdot, T) \in L_s^2$  then the error made by replacing  $\int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \dots dM$  by  $\int_{-\infty}^{\infty} \dots dM$  is  $\mathcal{O}(\varepsilon^{s-1/2})$ . Hence by equating the coefficients of  $\varepsilon^3$  and  $\varepsilon^5$  to zero we find  $c^2 = \lambda_1''(0)/2$  and  $\hat{A}$  to satisfy

$$-2ic\partial_T \hat{A}(K, T) - \lambda_1''''(0) K^3 \hat{A}(K, T) / 24 + \int_{-\infty}^{\infty} \nu_2 K \hat{A}(K-M, T) \hat{A}(M, T) dM = 0,$$

respectively,  $A$  to satisfy the KdV equation

$$(18) \quad 2c\partial_T A(X, T) + \lambda_1''''(0) \partial_X^3 A(X, T) / 24 + \nu_2 \partial_X (A(X, T)^2) = 0.$$

**3.2.2. The inviscid Burgers equation.** Due to the explanations in the Appendix A we restrict to the case  $\alpha = 1$ . We insert the inviscid Burgers approximation  $\varepsilon^\alpha \Psi$ , which is defined via (14) for  $\alpha = 1$ , into  $\text{Res}(\tilde{u})$ . We find with  $\tilde{u}_1(l, t) = \tilde{A}(K, T) \mathbf{E}$ ,  $\mathbf{E} = e^{i\varepsilon Kct}$ ,  $T = \varepsilon^2 t$ , and  $l = \varepsilon K$  that

$$\begin{aligned}
P_c(\text{Res}(\tilde{u}))(l, t) &= -\partial_t^2 \tilde{u}_1(l, t) - \lambda_1(l) \tilde{u}_1(l, t) \\
&\quad + \int_{-1/2}^{1/2} s_{11}^1(l, l-m, m) \tilde{u}_1(l-m, t) \tilde{u}_1(m, t) dm \\
&= \varepsilon^2 c^2 K^2 \tilde{A}(K, T) \mathbf{E} - 2\varepsilon^3 icK (\partial_T \tilde{A}(K, T)) \mathbf{E} - \varepsilon^4 (\partial_T^2 \tilde{A}(K, T)) \mathbf{E} \\
&\quad - \varepsilon^2 \lambda_1''(0) K^2 \tilde{A}(K, T) \mathbf{E} / 2 - \varepsilon^4 \lambda_1''''(0) K^4 \tilde{A}(K, T) \mathbf{E} / 24 + \mathcal{O}(\varepsilon^4) \\
&\quad + \varepsilon^3 \int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \nu_2 K^2 \tilde{A}(K-M, T) \tilde{A}(M, T) dM \mathbf{E} + \mathcal{O}(\varepsilon^4).
\end{aligned}$$

We proceed as above and equate the coefficients of  $\varepsilon^2$  and  $\varepsilon^3$  to zero. We find  $c^2 = \lambda_1''(0)/2$  and  $\hat{A}$  to satisfy

$$-2ic\partial_T \hat{A}(K, T) + \int_{-\infty}^{\infty} \nu_2 K \hat{A}(K-M, T) \hat{A}(M, T) dM = 0$$

respectively  $A$  to satisfy the inviscid Burgers equation

$$(19) \quad 2c\partial_T A(X, T) + \nu_2 \partial_X (A(X, T)^2) = 0.$$

**3.3. Derivation of the Whitham system.** The derivation of the Whitham system is much more involved since already in the derivation the  $\tilde{v}$  part has to be included. Due to the symmetry assumption **(SYM)** with  $u = u(x, t)$ , also  $u = u(-x, t)$  is a solution of (9). As a consequence in (9) all terms must contain an even number of  $\partial_x$ -derivatives. Since in Bloch space

$$\begin{aligned} u(-x, t) &= \int_{-1/2}^{1/2} \tilde{u}(-x, l) e^{-ilx} dl \\ &= - \int_{1/2}^{-1/2} \tilde{u}(-x, -l) e^{ilx} dl = \int_{-1/2}^{1/2} \tilde{u}(-x, -l) e^{ilx} dl \end{aligned}$$

with  $\tilde{u} = \tilde{u}(l, x, t)$ , also  $\tilde{u} = \tilde{u}(-l, -x, t)$  is a solution of the Bloch wave transformed system (13). As a consequence in (13) all terms must contain an even number of  $\partial_x$ -derivatives or  $il$ ,  $i(l-m)$ , or  $im$  factors, i.e., for instance  $il\partial_x$  can occur, but  $-l^2\partial_x$  not. Before we start with the derivation of the Whitham system we additionally need that in some of the kernel functions  $s_{j_1 j_2}^j$  at least one  $l$  factor occurs.

**Lemma 3.4.** *We have*

$$|s_{vv}^1(l, l-m, m)| \leq C|l|$$

and

$$|s_{11}^v(l, l-m, m)| \leq C(|l| + (l-m)^2 + m^2).$$

**Proof.** a) Using again the expansion (17) yields after some integration by parts that

$$\begin{aligned} & \int_0^{2\pi} \overline{\tilde{w}_1(l, x)} (\partial_x + il) (c(x) (\partial_x + il) \int_{-1/2}^{1/2} \tilde{v}(l-m, x, t) \tilde{v}(m, x, t) dm dx \\ &= \int_{-1/2}^{1/2} \int_0^{2\pi} c(x) (-il + il\partial_x g_1(x) + \mathcal{O}(l^2)) (\partial_x + il) (\tilde{v}(l-m, x, t) \tilde{v}(m, x, t)) dx dm \\ &= \mathcal{O}(l) \end{aligned}$$

b) As above we obtain

$$\begin{aligned} & s_{11}^v(l, l-m, m) \\ &= \int_0^{2\pi} \overline{\tilde{v}(l, x)} (\partial_x + il) (c(x) (\partial_x + il) (\tilde{w}_1(l-m, x) \tilde{w}_1(m, x))) dx \\ &= \int_0^{2\pi} \overline{\tilde{v}(l, x)} (\partial_x + il) (c(x) (\partial_x + il) \\ &\quad \times ((1 + i(l-m)g_1(x) + \mathcal{O}((l-m)^2))(1 + img_1(x) + \mathcal{O}(m^2)))) dx \\ &= \int_0^{2\pi} \overline{\tilde{v}(l, x)} (\partial_x + il) (c(x) (il + il\partial_x g_1(x) + \mathcal{O}((l-m)^2 + m^2))) dx \\ &= \mathcal{O}(|l| + (l-m)^2 + m^2). \end{aligned}$$

□

For the derivation of the Whitham system we make the ansatz

$$(20) \quad \tilde{u}_1(l, t) = \varepsilon^{-1} \tilde{A}(K, T) \quad \text{and} \quad \tilde{v}(l, x, t) = \tilde{B}(K, x, T).$$

where  $T = \varepsilon t$ , and  $l = \varepsilon K$ . With  $\tilde{u}(l, x, t) = \tilde{u}_1(l, t) \tilde{w}_1(l, x) + \tilde{v}(l, x, t)$  we find that

$$\begin{aligned} P_c(\text{Res}(\tilde{u}))(l, t) &= -\partial_t^2 \tilde{u}_1(l, t) - \lambda_1(l) \tilde{u}_1(l, t) + P_c(l) N_l(\partial_x)(\tilde{u})(l, t) \\ &= -\varepsilon \partial_T^2 \tilde{A}(K, T) - \varepsilon \lambda_1''(0) K^2 \tilde{A}(K, T)/2 + \mathcal{O}(\varepsilon^3) \\ &\quad + P_c(\varepsilon K) N_{\varepsilon K}(\partial_x)(\tilde{u})(\varepsilon K, T/\varepsilon) \end{aligned}$$

and

$$\begin{aligned} P_s(\text{Res}(\tilde{u}))(l, x, t) &= -\partial_t^2 \tilde{v}(l, x, t) - L_l(\partial_x) \tilde{v}(l, x, t) + P_s(l) N_l(\partial_x)(\tilde{u})(l, x, t) \\ &= -\varepsilon^2 \partial_T^2 \tilde{B}(K, x, T) - \tilde{L}_{\varepsilon K}(\partial_x) \tilde{B}(K, x, T) \\ &\quad + P_s(l) N_{\varepsilon K}(\partial_x)(\tilde{u})(\varepsilon K, x, T/\varepsilon). \end{aligned}$$

Since  $P_s(\varepsilon K) N_{\varepsilon K}(\partial_x)(\tilde{u})$  is quadratic w.r.t.  $\tilde{u}$  and since  $\tilde{L}_{\varepsilon K}$  is invertible on the range of  $P_s(\varepsilon K)$  we can use the implicit function theorem to solve

$$-\tilde{L}_{\varepsilon K}(\partial_x) \tilde{B}(K, x, T) + P_s(\varepsilon K) N_{\varepsilon K}(\partial_x)(\tilde{u})(\varepsilon K, x, T/\varepsilon) = 0$$

w.r.t.  $\tilde{B} = H(\tilde{A})(K, x, T)$  for sufficiently small  $\tilde{A}$ . Note that we kept our notation and still wrote  $T/\varepsilon$  in the arguments of  $N$  although in fact it only depends on  $T$ . We insert  $\tilde{B} = H(\tilde{A})(K, x, T)$  into the first equation and obtain

$$\begin{aligned} P_c(\text{Res}(\tilde{u}))(l, t) &= -\varepsilon \partial_T^2 \tilde{A}(K, T) - \varepsilon \lambda_1''(0) K^2 \tilde{A}(K, T)/2 + \mathcal{O}(\varepsilon^3) \\ &\quad + P_c(\varepsilon K) N_{\varepsilon K}(\partial_x)(\varepsilon^{-1} \tilde{A}(K, T) \tilde{w}_1(\varepsilon K, x) \\ &\quad + H(\tilde{A})(K, x, T))(\varepsilon K, T/\varepsilon) \end{aligned}$$

The Whitham system occurs by expanding the right hand side w.r.t.  $\varepsilon$  and by equating the coefficient in front of  $\varepsilon^1$  to zero. We obtain in a first step

$$\partial_T^2 \tilde{A}(K, T) + \lambda_1''(0) K^2 \tilde{A}(K, T)/2 + \tilde{G}(\tilde{A})(K, T) = 0$$

where  $\tilde{G}$  is a nonlinear function that can be written as

$$\tilde{G}(\tilde{A})(K, T) = -\chi_{[-\delta_0/4, \delta_0/4]}(\varepsilon K) \sum_{j=2}^{\infty} s_j i K \int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \tilde{A}^{*(j-1)}(K-M) i M \tilde{A}(M) dM$$

with coefficients  $s_j$ . The factor  $iK$  comes from Lemma 3.3 and Lemma 3.4 a), the factor  $iM$  from the fact that due to the reflection symmetry we need an even number of such factors and due to the long wave character of the approximation we have exactly two such factors at  $\varepsilon$ . Replacing via (15) the Bloch transform  $\tilde{A}(K, T)$  by the Fourier transform  $\hat{A}(K, T)$  finally gives Whitham's system

$$(21) \quad \partial_T^2 \hat{A}(K, T) + \lambda_1''(0) K^2 \hat{A}(K, T)/2 + \hat{G}(\hat{A})(K, T) = 0$$



in Fourier space where  $\widehat{G}$  is a nonlinear function that can be written as

$$\widehat{G}(\widehat{A})(K, T) = - \sum_{j=2}^{\infty} s_j i K \int_{-\infty}^{\infty} \widehat{A}^{*(j-1)}(K-M) i M \widehat{A}(M) dM.$$

In physical space we have

$$G(A)(X, T) = - \sum_{j=2}^{\infty} s_j \partial_X (A^{j-1} \partial_X A) = - \partial_X^2 \sum_{j=2}^{\infty} s_j A^j / j$$

such that Whitham's system finally can be written as

$$(22) \quad \partial_T^2 A = \partial_X^2 \mathcal{H}(A), \quad \text{with} \quad \mathcal{H}(A) = -\lambda_1''(0) A / 2 - \sum_{j=2}^{\infty} s_j A^j / j.$$

#### 4. ESTIMATES FOR THE RESIDUAL

After the derivation of the amplitude equations we estimate the so called residual, the terms which do not cancel after inserting the approximation into (9). In order to have estimates as in the spatially homogeneous case for the residual terms in terms of  $\varepsilon$  we have to modify our approximations with higher order terms.

**The improved KdV approximation.** For the construction of the improved KdV approximation we proceed as for the derivation of the Whitham system. With  $\mathbf{E} = e^{i\varepsilon Kct}$ ,  $T = \varepsilon^3 t$ , and  $l = \varepsilon K$  we make the ansatz

$$\begin{aligned} \widetilde{u}_1(l, t) &= \varepsilon \widetilde{A}(K, T) \mathbf{E} \\ \widetilde{v}(l, x, t) &= \varepsilon^4 \widetilde{B}(K, x, T) \mathbf{E} + \varepsilon^5 \widetilde{B}_2(K, x, T) \mathbf{E} + \varepsilon^3 \widetilde{B}_3(K, x, T) \mathbf{E}. \end{aligned}$$

With  $\widetilde{u}(l, x, t) = \widetilde{u}_1(l, t) \widetilde{w}_1(l, x) + \widetilde{v}(l, x, t)$ ,  $T = \varepsilon t$ , and  $l = \varepsilon K$  we find that

$$\begin{aligned} P_c(\text{Res}(\widetilde{u}))(l, t) &= -\partial_t^2 \widetilde{u}_1(l, t) - \lambda_1(l) \widetilde{u}_1(l, t) + P_c(l) N_l(\partial_x)(\widetilde{u})(l, t) \\ &= \varepsilon^3 c^2 K^2 \widetilde{A}(K, T) \mathbf{E} - 2\varepsilon^5 icK (\partial_T \widetilde{A}(K, T)) \mathbf{E} - \varepsilon^7 (\partial_T^2 \widetilde{A}(K, T)) \mathbf{E} \\ &\quad - \varepsilon^3 \lambda_1''(0) K^2 \widetilde{A}(K, T) \mathbf{E} / 2 - \varepsilon^5 \lambda_1''''(0) K^4 \widetilde{A}(K, T) \mathbf{E} / 24 + \mathcal{O}(\varepsilon^7) \\ &\quad + \varepsilon^5 \int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \nu_2 K^2 \widetilde{A}(K-M, T) \widetilde{A}(M, T) dM \mathbf{E} + \mathcal{O}(\varepsilon^7) = \mathcal{O}(\varepsilon^7) \end{aligned}$$

if we choose  $c$  and  $\widetilde{A}$  as above. We have  $\mathcal{O}(\varepsilon^7)$  and not  $\mathcal{O}(\varepsilon^6)$  since  $P_c(\text{Res}(\widetilde{u}))(l, t)$  does not depend on  $x$  and has to be even w.r.t. factors in  $l$ , i.e.,  $\varepsilon^5 K^4 \widetilde{A}(K, T)$  is

allowed, but not  $\varepsilon^6 K^5 \tilde{A}(K, T)$ . Next we have

$$\begin{aligned}
P_s(\text{Res}(\tilde{u}))(l, x, t) &= -\partial_t^2 \tilde{v}(l, x, t) - L_l(\partial_x) \tilde{v}(l, x, t) + P_s(l) N_l(\partial_x)(\tilde{u})(l, x, t) \\
&= c^2 K^2 (\varepsilon^6 \tilde{B}(K, x, T) + \varepsilon^7 \tilde{B}_2(K, x, T) + \varepsilon^8 \tilde{B}_3(K, x, T)) \mathbf{E} \\
&\quad - 2icK (\varepsilon^8 \partial_T \tilde{B}(K, x, T) + \varepsilon^9 \partial_T \tilde{B}_2(K, x, T) + \varepsilon^{10} \partial_T \tilde{B}_3(K, x, T)) \mathbf{E} \\
&\quad - (\varepsilon^{10} \partial_T^2 \tilde{B}(K, x, T) + \varepsilon^{11} \partial_T^2 \tilde{B}_2(K, x, T) + \varepsilon^{12} \partial_T^2 \tilde{B}_3(K, x, T)) \mathbf{E} \\
&\quad - (\varepsilon^4 \tilde{L}_{\varepsilon K}(\partial_x) \tilde{B}(K, x, T) + \varepsilon^5 \tilde{L}_{\varepsilon K}(\partial_x) \tilde{B}_2(K, x, T) \\
&\quad + \varepsilon^6 \tilde{L}_{\varepsilon K}(\partial_x) \tilde{B}_3(K, x, T)) \mathbf{E} + P_s(\varepsilon K) N_{\varepsilon K}(\partial_x)(\tilde{u})(\varepsilon K, x, T/\varepsilon)
\end{aligned}$$

where we expand

$$P_s(l) N_{\varepsilon K}(\partial_x)(\tilde{u})(\varepsilon K, x, T/\varepsilon) = (\varepsilon^4 F_4(\tilde{A}) + \varepsilon^5 F_5(\tilde{A}) + \varepsilon^6 F_6(\tilde{A}, \tilde{B}) + \mathcal{O}(\varepsilon^7))(K, x, T) \mathbf{E}.$$

If we set

$$\begin{aligned}
0 &= -\tilde{L}_{\varepsilon K}(\partial_x) \tilde{B}(K, x, T) + \varepsilon^4 F_4(\tilde{A})(K, x, T), \\
0 &= -\tilde{L}_{\varepsilon K}(\partial_x) \tilde{B}_2(K, x, T) + \varepsilon^4 F_5(\tilde{A})(K, x, T), \\
0 &= -\tilde{L}_{\varepsilon K}(\partial_x) \tilde{B}_3(K, x, T) + \varepsilon^4 F_6(\tilde{A}, \tilde{B})(K, x, T) + c^2 K^2 \tilde{B}(K, x, T),
\end{aligned}$$

we finally have

$$P_s(\text{Res}(\tilde{u}))(l, x, t) = \mathcal{O}(\varepsilon^7).$$

The functions  $\tilde{B}$ ,  $\tilde{B}_2$ , and  $\tilde{B}_3$  are well-defined since  $\tilde{L}_{\varepsilon K}$  can be inverted on the range of  $P_s(\varepsilon K)$ .

**The improved inciscid Burgers approximation.** We leave this part to the reader. We refer to Appendix A where the modified approximation is discussed for the spatially homogeneous situation.

**The improved Whitham approximation.** We need the residual formally to be of order  $\mathcal{O}(\varepsilon^3)$ . With the previous approximation we already have  $\mathcal{O}(\varepsilon^3)$  for the  $P_c$ -part of the residual again due to symmetry reasons, but we only have  $\mathcal{O}(\varepsilon^2)$  for the  $P_s$ -part. As above we modify our ansatz into

$$\tilde{u}_1(l, t) = \varepsilon^{-1} \tilde{A}(K, T) \quad \text{and} \quad \tilde{v}(l, x, t) = \tilde{B}(K, x, T) + \varepsilon^2 \tilde{B}_2(K, x, T).$$

We define  $\tilde{A}$  and  $\tilde{B}$  exactly as above and  $\tilde{B}_2$  as solution of

$$-\partial_T^2 \tilde{B}(K, x, T) - \tilde{L}_{\varepsilon K}(\partial_x) \tilde{B}_2(K, x, T) = 0$$

which is again well-defined due the fact that  $\tilde{L}_{\varepsilon K}$  can be inverted on the range of  $P_s(\varepsilon K)$ .

For all three approximations we gain a factor  $\varepsilon^{1/2}$  when we estimate the error in  $L^2$ -based spaces due to the scaling properties of the  $L^2$  norm. Since the error

made by the various approximations will be estimated in physical space via energy estimates we conclude for the KdV approximation, for the inviscid Burgers approximation, and for the Whitham approximation that

**Lemma 4.1.** *Let  $A \in C([0, T_0], H^6)$  be a solution of the KdV equation (18). Then there exist  $\varepsilon_0 > 0$ ,  $C_{\text{res}}$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have*

$$\sup_{t \in [0, T_0/\varepsilon^3]} \|\text{Res}(\varepsilon^2 \Psi(\cdot, t, \varepsilon))\|_{H^1} \leq C_{\text{res}} \varepsilon^{15/2}$$

and

$$\sup_{t \in [0, T_0/\varepsilon^3]} \|\partial_x^{-1} \text{Res}(\varepsilon^2 \Psi(\cdot, t, \varepsilon))\|_{H^1} \leq C_{\text{res}} \varepsilon^{13/2}.$$

**Lemma 4.2.** *Let  $\alpha = 1$  and let  $A \in C([0, T_0], H^4)$  be a solution of the inviscid Burgers equation (3). Then there exist  $\varepsilon_0 > 0$ ,  $C_{\text{res}}$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have*

$$\sup_{t \in [0, T_0/\varepsilon^{1+\alpha}]} \|\text{Res}(\varepsilon^\alpha \Psi(\cdot, t, \varepsilon))\|_{H^1} \leq C_{\text{res}} \varepsilon^{(7+4\alpha)/2}$$

and

$$\sup_{t \in [0, T_0/\varepsilon^{1+\alpha}]} \|\partial_x^{-1} \text{Res}(\varepsilon^\alpha \Psi(\cdot, t, \varepsilon))\|_{H^1} \leq C_{\text{res}} \varepsilon^{(5+4\alpha)/2}.$$

**Lemma 4.3.** *Let  $A \in C([0, T_0], H^4)$  be a solution of the Whitham equation (7). Then there exist  $\varepsilon_0 > 0$ ,  $C_{\text{res}}$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have*

$$\sup_{t \in [0, T_0/\varepsilon]} \|\text{Res}(\Psi(\cdot, t, \varepsilon))\|_{H^1} \leq C_{\text{res}} \varepsilon^{7/2}$$

and

$$\sup_{t \in [0, T_0/\varepsilon]} \|\partial_x^{-1} \text{Res}(\Psi(\cdot, t, \varepsilon))\|_{H^1} \leq C_{\text{res}} \varepsilon^{5/2}.$$

## 5. THE ERROR ESTIMATES

As for spatially homogeneous case the proofs given for the KdV approximations transfer more or less line for line into proofs for the justification of the inviscid Burgers equation and of the Whitham system. Our approximation results are as follows

**Theorem 5.1.** *Let  $A \in C([0, T_0], H^6(\mathbb{R}))$  be a solution of the KdV equation (18). Then there exist  $\varepsilon_0 > 0$ ,  $C > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have solutions  $u \in C([0, T_0/\varepsilon^3], H^2)$  of the spatially periodic Boussinesq model (9) with*

$$\sup_{t \in [0, T_0/\varepsilon^3]} \|u(\cdot, t) - \varepsilon^2 A(\varepsilon(\cdot - t), \varepsilon^3 t)\|_{H^2} \leq C \varepsilon^{5/2}.$$

**Theorem 5.2.** *Let  $\alpha = 1$  and let  $A \in C([0, T_0], H^4(\mathbb{R}))$  be a solution of the inviscid Burgers equation (18). Then there exist  $\varepsilon_0 > 0$ ,  $C > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have solutions  $u \in C([0, T_0/\varepsilon^3], H^2)$  of the spatially periodic Boussinesq model (9) with*

$$\sup_{t \in [0, T_0/\varepsilon^{1+\alpha}]} \|u(\cdot, t) - \varepsilon^\alpha A(\varepsilon(\cdot - t), \varepsilon^{1+\alpha}t)\|_{H^2} \leq C\varepsilon^{(1+2\alpha)/2}.$$

**Theorem 5.3.** *There exists a  $C_1 > 0$  such that the following holds. Let  $U \in C([0, T_0], H^4)$  be a solution of the Whitham system (7) with*

$$\sup_{T \in [0, T_0]} \|U(\cdot, T)\|_{H^4} \leq C_1.$$

*Then there exist  $\varepsilon_0 > 0$  and  $C > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have solutions  $u \in C^0([0, T_0/\varepsilon^3], H^2)$  of our spatially periodic Boussinesq model (9), such that*

$$\sup_{t \in [0, T_0/\varepsilon^3]} \|u(\cdot, t) - U(\varepsilon\cdot, \varepsilon t)\|_{H^2} \leq C_2\varepsilon^{1/2}.$$

**Proof of the Theorems 5.1-5.3.** Since we already have the estimates for the residuals in the Lemmas 4.1-4.3 from this point on the remaining estimates can be handled exactly the same. The case  $\alpha = 0$  corresponds to the Whitham approximation and the case  $\alpha = 2$  to the KdV approximation.

The difference  $\varepsilon^{(3+2\alpha)/2}R = u - \varepsilon^\alpha\Psi$  satisfies

$$(23) \quad \begin{aligned} \partial_t^2 R &= \partial_x(a\partial_x R) - \partial_x^2(b\partial_x^2 R) + 2\partial_x(c\partial_x(\varepsilon^\alpha\Psi R)) \\ &\quad + \varepsilon^{(3+2\alpha)/2}\partial_x(c\partial_x(R^2)) + \varepsilon^{-(3+2\alpha)/2}\text{Res}(\varepsilon^\alpha\Psi). \end{aligned}$$

The first three terms on the right hand side can be written as

$$\partial_x(a\partial_x R) - \partial_x^2(b\partial_x^2 R) + 2\partial_x(c\varepsilon^\alpha\Psi\partial_x R) + 2\partial_x(c(\partial_x\varepsilon^\alpha\Psi)R)$$

The last term is of order  $\mathcal{O}(\varepsilon^{1+\alpha})$  due to the long wave character of the approximation  $\varepsilon^\alpha\Psi$ . More essential the first three terms can be written as  $\partial_x(B(\partial_x R))$  where  $B$  is the self-adjoint operator

$$B = (a + 2c\varepsilon^\alpha\Psi) - \partial_x(b\partial_x).$$

In case  $\alpha > 0$  for sufficiently small  $\varepsilon > 0$  and in case  $\alpha = 0$  for sufficiently small  $\|\Psi\|_{C_b^0}$  the linear operator  $B$  is positive definite. Hence there exists a positive-definite self-adjoint operator  $\mathcal{A}$  with  $\mathcal{A}^2 = B$ . The associated operator norm  $\|\cdot\|_{\mathcal{A}} = \|\mathcal{A}\cdot\|_{L^2}$  is then equivalent to the  $H^1$ -norm and  $\mathcal{A}^{-1}$  is a bounded operator from  $L^2$  to  $H^1$ . Hence the equation for the error can be written as

$$(24) \quad \begin{aligned} \partial_t^2 R &= \partial_x(\mathcal{A}^2(\partial_x R)) + 2\partial_x(c(\partial_x\varepsilon^\alpha\Psi)R) \\ &\quad + \varepsilon^{(3+2\alpha)/2}\partial_x(c\partial_x(R^2)) + \varepsilon^{-(3+2\alpha)/2}\text{Res}(\varepsilon^\alpha\Psi). \end{aligned}$$

In order to bound the solutions of (24) we use energy estimates. Therefore, we first multiply (24) with  $\partial_t R$  and integrate the obtained expression w.r.t.  $x$ . We

obtain

$$\int (\partial_t R) \partial_t^2 R dx = \partial_t \int (\partial_t R)^2 dx / 2$$

and

$$\begin{aligned} & \int (\partial_t R) \partial_x (\mathcal{A}^2 (\partial_x R)) dx \\ = & - \int (\partial_t \partial_x R) (\mathcal{A}^2 (\partial_x R)) dx = - \int (\mathcal{A} \partial_t \partial_x R) (\mathcal{A} \partial_x R) dx \\ = & - \int (\partial_t (\mathcal{A} \partial_x R)) (\mathcal{A} \partial_x R) dx - \int ([\partial_t, \mathcal{A}] \partial_x R) (\mathcal{A} \partial_x R) dx \\ = & - \partial_t \int (\mathcal{A} \partial_x R)^2 dx / 2 - \int ([\partial_t, \mathcal{A}] \partial_x R) (\mathcal{A} \partial_x R) dx. \end{aligned}$$

where

$$[\partial_t, \mathcal{A}] \cdot = \partial_t (\mathcal{A} \cdot) - \mathcal{A} \partial_t \cdot$$

is the commutator of the operators  $\mathcal{A}$  and  $\partial_t$ . Moreover, we estimate

$$\begin{aligned} \left| \int (\partial_t R) 2 \partial_x (c (\partial_x \varepsilon^\alpha \Psi) R) dx \right| & \leq C \varepsilon^{1+\alpha} \|\partial_t R\|_{L^2} \|R\|_{H^1}, \\ \left| \int (\partial_t R) \varepsilon^{(3+2\alpha)/2} \partial_x (c \partial_x (R^2)) dx \right| & \leq C \varepsilon^{(3+2\alpha)/2} \|\partial_t R\|_{L^2} \|R\|_{H^2}^2, \\ \left| \int (\partial_t R) \varepsilon^{-(3+2\alpha)/2} \text{Res}(\varepsilon^\alpha \Psi) dx \right| & \leq C \varepsilon^{2+\alpha} \|\partial_t R\|_{L^2} \end{aligned}$$

where we used the Lemmas 4.1-4.3. Finally we have

$$[\partial_t, \mathcal{A}] \partial_x R = (\partial_t \mathcal{A}) \partial_x R$$

such that

$$\int ([\partial_t, \mathcal{A}] \partial_x R) (\mathcal{A} \partial_x R) dx = \int ((\partial_t \mathcal{A}) \partial_x R) (\mathcal{A} \partial_x R) dx.$$

In order to control this term we first note that

$$(\partial_t \mathcal{A}) \mathcal{A} + \mathcal{A} \partial_t \mathcal{A} = \partial_t (\mathcal{A}^2) = 2c \partial_t (\varepsilon^\alpha \Psi)$$

and

$$((\partial_t \mathcal{A}) u, v)_{L^2} = (u, (\partial_t \mathcal{A}) v)_{L^2}$$

which follows from differentiating the associated formula for  $\mathcal{A}$  w.r.t.  $t$  such that

$$\begin{aligned}
|2 \int ((\partial_t \mathcal{A}) \partial_x R) (\mathcal{A} \partial_x R) dx| &= | \int (\mathcal{A} (\partial_t \mathcal{A}) \partial_x R) \partial_x R + \partial_x R (\partial_t \mathcal{A} (\mathcal{A} \partial_x R)) dx | \\
&= | \int \partial_x R (\mathcal{A} \partial_t \mathcal{A} + (\partial_t \mathcal{A}) \mathcal{A}) \partial_x R dx | \\
&= | \int 2c(\partial_t(\varepsilon^\alpha \Psi)) (\partial_x R)^2 dx | \\
&\leq 2 \sup_{x \in \mathbb{R}} |c(x) \partial_t(\varepsilon^\alpha \Psi(x, t))| \|\partial_x R\|_{L^2}^2 = \mathcal{O}(\varepsilon^{1+\alpha}) \|\partial_x R\|_{L^2}^2.
\end{aligned}$$

In order to get a bound for the  $L^2$ -norm of  $R$  and not only of its derivatives we secondly multiply " $\partial_x^{-1}(24)$ " with  $\mathcal{A}^{-2} \partial_x^{-1} \partial_t R$  and integrate the expression obtained in this way w.r.t.  $x$ . We find

$$\begin{aligned}
\int (\mathcal{A}^{-2} \partial_x^{-1} \partial_t R) \partial_x^{-1} \partial_t^2 R dx &= \int (\mathcal{A}^{-1} \partial_x^{-1} \partial_t R) \mathcal{A}^{-1} \partial_t \partial_x^{-1} \partial_t R dx \\
&= \partial_t \int (\mathcal{A}^{-1} \partial_x^{-1} \partial_t R)^2 dx / 2 \\
&\quad - \int (\mathcal{A}^{-1} \partial_x^{-1} \partial_t R) [\partial_t, \mathcal{A}^{-1}] \partial_x^{-1} \partial_t R dx, \\
\int (\mathcal{A}^{-2} \partial_x^{-1} \partial_t R) \partial_x^{-1} \partial_x \mathcal{A}^2 \partial_x R dx &= -\partial_t \int R^2 dx / 2.
\end{aligned}$$

Moreover, using  $\mathcal{A}^{-1} : L^2 \rightarrow H^1$  and the self-adjointness of  $\mathcal{A}^{-1}$  we estimate

$$\begin{aligned}
| \int (\mathcal{A}^{-2} \partial_x^{-1} \partial_t R) 2 \partial_x^{-1} \partial_x (c(\partial_x \varepsilon^\alpha \Psi) R) dx | &= | \int (\mathcal{A}^{-1} \partial_x^{-1} \partial_t R) 2 \mathcal{A}^{-1} (c(\partial_x \varepsilon^\alpha \Psi) R) dx | \\
&\leq C \varepsilon^{1+\alpha} \|\partial_x^{-1} \partial_t R\|_{L^2} \|R\|_{L^2}, \\
| \int (\mathcal{A}^{-2} \partial_x^{-1} \partial_t R) \varepsilon^{(3+2\alpha)/2} \partial_x^{-1} \partial_x (c \partial_x (R^2)) dx | &= | \int (\mathcal{A}^{-1} \partial_x^{-1} \partial_t R) \varepsilon^{(3+2\alpha)/2} \mathcal{A}^{-1} (c \partial_x (R^2)) dx | \\
&\leq C \varepsilon^{(3+2\alpha)/2} \|\partial_x^{-1} \partial_t R\|_{L^2} \|R\|_{H^1}^2, \\
| \int (\mathcal{A}^{-2} \partial_x^{-1} \partial_t R) \varepsilon^{-(3+2\alpha)/2} \partial_x^{-1} \text{Res}(\varepsilon^\alpha \Psi) dx | &\leq C \varepsilon^{1+\alpha} \|\partial_x^{-1} \partial_t R\|_{L^2}.
\end{aligned}$$

where we used again the Lemmas 4.1-4.3. Finally we have

$$[\partial_t, \mathcal{A}^{-1}] \partial_x R = (\partial_t \mathcal{A}^{-1}) \partial_x R$$

such that

$$\int (\mathcal{A}^{-1} \partial_x^{-1} \partial_t R) [\partial_t, \mathcal{A}^{-1}] \partial_x^{-1} \partial_t R dx = \int (\mathcal{A}^{-1} \partial_x^{-1} \partial_t R) (\partial_t \mathcal{A}^{-1}) \partial_x^{-1} \partial_t R dx.$$

We write this as half of

$$\begin{aligned}
& \int ((\partial_t \mathcal{A}^{-1}) \partial_x^{-1} \partial_t R) (\mathcal{A}^{-1} \partial_x^{-1} \partial_t R) dx + \int (\mathcal{A}^{-1} \partial_x^{-1} \partial_t R) (\partial_t \mathcal{A}^{-1}) \partial_x^{-1} \partial_t R dx \\
&= \int (\partial_x^{-1} \partial_t R) ((\partial_t \mathcal{A}^{-1}) \mathcal{A}^{-1} + \mathcal{A}^{-1} \partial_t \mathcal{A}^{-1}) \partial_x^{-1} \partial_t R dx \\
&= \int (\partial_x^{-1} \partial_t R) (\partial_t (\mathcal{A}^{-2})) \partial_x^{-1} \partial_t R dx =: s_1
\end{aligned}$$

From

$$\partial_t (\mathcal{A}^2 \mathcal{A}^{-2}) = (\partial_t (\mathcal{A}^2)) \mathcal{A}^{-2} + \mathcal{A}^2 \partial_t (\mathcal{A}^{-2}) = 0$$

it follows that

$$\partial_t (\mathcal{A}^{-2}) = -\mathcal{A}^{-2} (\partial_t (\mathcal{A}^2)) \mathcal{A}^{-2} = -\mathcal{A}^{-2} (\partial_t (\varepsilon^\alpha \Psi)) \mathcal{A}^{-2}$$

such that

$$s_1 = \int (\mathcal{A}^{-2} \partial_x^{-1} \partial_t R) (\partial_t (\varepsilon^\alpha \Psi)) (\mathcal{A}^{-2} \partial_x^{-1} \partial_t R) dx = \mathcal{O}(\varepsilon^{1+\alpha}) \|\mathcal{A}^{-2} \partial_x^{-1} \partial_t R\|_{L^2}^2$$

which can be bounded by  $\mathcal{O}(\varepsilon^{1+\alpha}) \|\mathcal{A}^{-1} \partial_x^{-1} \partial_t R\|_{L^2}^2$ . If we define

$$E(t) = \frac{1}{2} (\|\partial_t R\|_{L^2}^2 + \|\mathcal{A}^{-1} \partial_x^{-1} \partial_t R\|_{L^2}^2 + \|R\|_{L^2}^2 + \|\mathcal{A} \partial_x R\|_{L^2}^2).$$

we find

$$\begin{aligned}
\frac{d}{dt} E &\leq C_1 \varepsilon^{1+\alpha} E + C_2 \varepsilon^{(3+2\alpha)/2} E^{3/2} + C_3 \varepsilon^{1+\alpha} E^{1/2} \\
&\leq C_1 \varepsilon^{1+\alpha} E + C_2 \varepsilon^{(3+2\alpha)/2} E^{3/2} + C_3 \varepsilon^{1+\alpha} + C_3 \varepsilon^{1+\alpha} E,
\end{aligned}$$

with constants  $C_1$ ,  $C_2$ , and  $C_3$  independent of  $0 < \varepsilon \ll 1$  since all the  $\|\partial_t R\|_{L^2}$ ,  $\|\mathcal{A}^{-1} \partial_x^{-1} \partial_t R\|_{L^2}$ , etc. appearing above can be estimated by  $E^{1/2}$ . Choosing  $\varepsilon^{1/2} E^{1/2} \leq 1$  gives

$$\frac{d}{dt} E(t) \leq (C_1 + C_2 + C_3) \varepsilon^{1+\alpha} E + C_3 \varepsilon^{1+\alpha}$$

which can be estimated with Gronwall's inequality and yields

$$E(t) \leq C_3 T_0 e^{(C_1 + C_2 + C_3) T_0} =: M$$

for all  $0 \leq \varepsilon^{1+\alpha} t \leq T_0$ . Choosing  $\varepsilon_0 > 0$  so small that  $\varepsilon_0^{1/2} M^{1/2} \leq 1$  gives the required estimate first for  $E(t)$ . Since in case  $\alpha > 0$  for sufficiently small  $\varepsilon > 0$  and case  $\alpha = 0$  for sufficiently small  $\|\Psi\|_{C_b^2}$  the quantity  $E^{1/2}$  equivalent to the  $H^2$ -norm of  $R$  we are done with the proof of the Theorems 5.1-5.3.  $\square$

## 6. DISCUSSION

It is the purpose of this section to give some heuristic arguments why the previous approach works and to put the approach in some larger framework.

The error equation (10) to the spatially homogeneous Boussinesq equation (1) can be written in lowest order in the form of a Hamiltonian system, namely

$$\partial_t \begin{pmatrix} R \\ w \end{pmatrix} = \begin{pmatrix} \partial_x^2 R - \partial_x^4 R + \varepsilon^\alpha \Psi \partial_x^2 R + \mathcal{O}(\varepsilon^{1+\alpha}) \\ w \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \partial_R H \\ \partial_w H \end{pmatrix},$$

with the Hamiltonian

$$H = \frac{1}{2} \int w^2 + (\partial_x R)^2 + (\partial_x^2 R)^2 + \varepsilon^\alpha \Psi (\partial_x R)^2 dx$$

where for this presentation we used  $\partial_x \Psi = \mathcal{O}(\varepsilon)$ . This Hamiltonian is a part of our energy and it can be used to estimate parts of the  $H^2$  norm. Since  $\Psi$  depends on  $t$  the Hamiltonian is not conserved, but we have

$$(25) \quad \frac{d}{dt} H = \nabla H \cdot \partial_t \begin{pmatrix} R \\ w \end{pmatrix} + \partial_t H = 0 + \mathcal{O}(\varepsilon^{1+\alpha})$$

since  $\partial_t \Psi = \mathcal{O}(\varepsilon)$  due to the long wave character of the approximation.

In a similar way the spatially periodic case can be understood. The error equation (24) to the spatially homogeneous Boussinesq equation (9) can be written in lowest order in the form of a Hamiltonian system, namely

$$\partial_t \begin{pmatrix} R \\ w \end{pmatrix} = \begin{pmatrix} \partial_x (\mathcal{A}^2 (\partial_x R)) + \mathcal{O}(\varepsilon^{1+\alpha}) \\ w \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \partial_R H \\ \partial_w H \end{pmatrix} + \mathcal{O}(\varepsilon^{1+\alpha}),$$

with the Hamiltonian

$$H = \frac{1}{2} \int w^2 + (\mathcal{A} \partial_x R)^2 dx$$

where for this presentation we used  $\partial_x \Psi = \mathcal{O}(\varepsilon)$ . This Hamiltonian is a part of our energy and it can be used to estimate parts of the  $H^2$  norm. Since  $\mathcal{A}$  depends via  $\Psi$  on  $t$  the Hamiltonian is not conserved, but again we have (25) since  $\partial_t \Psi = \mathcal{O}(\varepsilon)$  due to the long wave character of the approximation.

As already said the paper was originally intended as the next step in generalizing a method which has been developed in [7] for the justification of the KdV approximation in situations when the KdV modes are resonant to other long wave modes respectively in [15] for the justification of the Whitham approximation. The normal form transforms which were used in the proofs of [7, 15] leave the energy surfaces invariant and can therefore be avoided by our 'good' choice of energy. Hence also the toy problem considered in [7, 15] can be handled with the presented approach if the nonlinear terms are modified in such a way that a Hamiltonian structure is observed.



## APPENDIX A. THE INVISCID BURGERS APPROXIMATION

It is the goal of this appendix to provide more details about the derivation and the justification via error estimates for the inviscid Burgers approximation. Inserting the ansatz

$$\varepsilon^\alpha \Psi(x, t) = \varepsilon^\alpha A(\varepsilon(x - t), \varepsilon^{1+\alpha} t)$$

with  $\alpha \in (0, 2)$  into the homogeneous Boussinesq equation (9) gives the residual

$$\begin{aligned} \text{Res}(u)(x, t) &= -\partial_t^2 u(x, t) + \partial_x^2 u(x, t) - \partial_x^4 u(x, t) + \partial_x^2 (u(x, t)^2) \\ &= \varepsilon^{\alpha+4} \partial_X^4 A + \varepsilon^{3\alpha+2} \partial_T^2 A \end{aligned}$$

and  $A$  to satisfy the inviscid Burgers equation

$$\partial_T A = -\frac{1}{2} \partial_X (A^2)$$

if the coefficient of  $\varepsilon^{2\alpha+2}$  is put to zero. However, the residual is too large for the analysis made in Section 2. By adding higher order terms to the approximation we obtain the estimates stated in Remark 2.3, namely

$$\|\text{Res}(\varepsilon^\alpha \Psi(\cdot, t, \varepsilon))\|_{L^2} = \mathcal{O}(\varepsilon^{(7+4\alpha)/2}) \quad \text{and} \quad \|\partial_x^{-1} \text{Res}(\varepsilon^\alpha \Psi(\cdot, t, \varepsilon))\|_{L^2} = \mathcal{O}(\varepsilon^{(5+4\alpha)/2}).$$

We consider the improved approximation

$$\varepsilon^\alpha \Psi(x, t) = \varepsilon^\alpha A(\varepsilon(x - t), \varepsilon^{1+\alpha} t) + \varepsilon^\beta B(\varepsilon(x - t), \varepsilon^{1+\alpha} t)$$

with  $\beta = \min\{2\alpha, 2\}$ . For the residual we find

$$\begin{aligned} \text{Res}(\varepsilon^2 \Psi) &= -2\varepsilon^{2+\alpha+\beta} \partial_T \partial_X B - \varepsilon^{2+2\alpha+\beta} \partial_T^2 B - \varepsilon^{4+\beta} \partial_X^4 B + 2\varepsilon^{2+\alpha+\beta} \partial_X^2 (AB) \\ &\quad + \varepsilon^{2+2\beta} \partial_X^2 (B^2) + \varepsilon^{\alpha+4} \partial_X^4 A + \varepsilon^{3\alpha+2} \partial_T^2 A. \end{aligned}$$

We choose  $B$  to satisfy

$$2\partial_T B = 2\partial_X (AB) + g$$

where

$$g = \begin{cases} \partial_X^{-1} \partial_T^2 A, & \text{for } \alpha \in (0, 1), \\ \partial_X^{-1} \partial_T^2 A + \partial_X^3 A, & \text{for } \alpha = 1, \\ \partial_X^3 A, & \text{for } \alpha \in (1, 2). \end{cases}$$

By this choice we have

$$|\text{Res}(\varepsilon^2 \Psi)| = \mathcal{O}(\max\{\chi_{\alpha \neq 1}(\alpha) \min\{\varepsilon^{\alpha+4}, \varepsilon^{3\alpha+2}\}, \varepsilon^{2+2\alpha+\beta}, \varepsilon^{4+\beta}, \varepsilon^{2+2\beta}\}).$$

Hence only for  $\alpha = 1$ , where  $\beta = 2$ , this is of order  $\mathcal{O}(\varepsilon^{4+2\alpha})$  which is the formal order which is necessary to obtain the  $L^2$  bound. For all other values of  $\alpha$  more additional terms are necessary. For  $\alpha \rightarrow 0$  and  $\alpha \rightarrow 2$  the number of such terms goes to infinity and more and more regularity is necessary. We refrain from discussing the solvability of this system of amplitude equations. This question is non-trivial since already for  $\alpha = 1$  the term  $\partial_X^{-1} \partial_T^2 A$  has to be computed which is

possible due to the fact that the temporal derivatives can be expressed as spatial derivatives via the inviscid Burgers equation, namely

$$\partial_T^2 A = -\frac{1}{2}\partial_T\partial_X(A^2) = -\partial_X(A\partial_TA) = \frac{1}{2}\partial_X(A\partial_X(A^2)) = \frac{1}{3}\partial_X^2(A^3).$$

Due to this presentation also the estimate for  $\partial_x^{-1}\text{Res}(\varepsilon^\alpha\Psi)$  can be obtained since now also  $\partial_T^2 B$  can be expressed as spatial derivatives.

## APPENDIX B. HIGHER REGULARITY RESULTS

It is the purpose of this section to explain how the approximation results can be transferred from  $H^2$  to  $H^m$  with  $m \geq 2$ . Due to the  $x$ -dependent coefficients energy estimates for the spatial derivatives turn out to be rather complicated. However, by considering time derivatives the previous ideas and energies still can be used. The spatial derivatives then can be estimated via the equation for the error, namely

$$(26) \quad \begin{aligned} LR &= \partial_t^2 R - 2\partial_x(c\partial_x(\varepsilon^\alpha\Psi R)) - R \\ &\quad - \varepsilon^{(3+2\alpha)/2}\partial_x(c\partial_x(R^2)) - \varepsilon^{-(3+2\alpha)/2}\text{Res}(\varepsilon^2\Psi). \end{aligned}$$

where

$$LR = \partial_x(a\partial_x R) - \partial_x^2(b\partial_x^2 R) - R.$$

The operator  $L$  is invertible and maps  $H^s$  into  $H^{s+4}$ , respectively  $C^m([0, T_0/\varepsilon^{1+\alpha}], H^s)$  into  $C^m([0, T_0/\varepsilon^{1+\alpha}], H^{s+4})$ . For  $R \in C^m([0, T_0/\varepsilon^{1+\alpha}], H^s)$  the right-hand side of (26) is in

$$C^{m-2}([0, T_0/\varepsilon^{1+\alpha}], H^s) \cap C^m([0, T_0/\varepsilon^{1+\alpha}], H^{s-2}).$$

An application of  $L^{-1}$  to (26) shows that

$$R \in C^{m-2}([0, T_0/\varepsilon^{1+\alpha}], H^{s+4}) \cap C^m([0, T_0/\varepsilon^{1+\alpha}], H^{s+2}).$$

Iterating this process shows that temporal derivatives can be transformed into spatial derivatives.

It remains to obtain the estimates for the temporal derivatives. In order to do so we differentiate the equation for the error  $m$  times w.r.t.  $t$ . We obtain an equation of the form

$$(27) \quad \partial_t^2(\partial_t^m R) = \partial_x(\mathcal{A}^2(\partial_x(\partial_t^m R))) + 2\partial_x(c(\partial_x\varepsilon^\alpha\Psi)(\partial_t^m R)) + \mathcal{O}(\varepsilon^{1+\alpha})$$

due to the fact that whenever a time derivative falls on  $\mathcal{A}$  or  $\Psi$  another  $\varepsilon$  is gained. In order to bound the solutions of (27) we use energy estimates. Therefore, we first multiply (27) with  $\partial_t^{m+1}R$  and integrate the obtained expression w.r.t.  $x$ . Next as above we multiply " $\partial_x^{-1}(27)$ " with  $\mathcal{A}^{-2}\partial_x^{-1}\partial_t R$  and integrate the expression obtained in this way w.r.t.  $x$ .

If we define

$$E_m(t) = \frac{1}{2} \left( \|\partial_t^{m+1}R\|_{L^2}^2 + \|\mathcal{A}^{-1}\partial_x^{-1}\partial_t^{m+1}R\|_{L^2}^2 + \|\partial_t^m R\|_{L^2}^2 + \|\mathcal{A}\partial_x\partial_t^m R\|_{L^2}^2 \right).$$

we find

$$\frac{d}{dt}E_m \leq C_1\varepsilon^{1+\alpha}E_m + C_2\varepsilon^{(3+2\alpha)/2}\mathcal{E}_m^{3/2} + C_3\varepsilon^{1+\alpha},$$

with constants  $C_1$ ,  $C_2$ , and  $C_3$  independent of  $0 < \varepsilon \ll 1$  and  $\mathcal{E}_m = E + \dots + E_m$ . Summing up all estimates for the  $E_j$  for  $j = 0, \dots, m$  yields a similar inequality for  $\mathcal{E}_m$ . Applying Gronwall's inequality to this inequality gives for instance

**Theorem B.1.** *Fix  $s \in \mathbb{N}$  and let  $A \in C([0, T_0], H^{6+s}(\mathbb{R}))$  be a solution of the KdV equation (18). Then there exist  $\varepsilon_0 > 0$ ,  $C > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0)$  we have solutions  $u \in C([0, T_0/\varepsilon^3], H^{2+s})$  of the spatially periodic Boussinesq model (9) with*

$$\sup_{t \in [0, T_0/\varepsilon^3]} \|u(\cdot, t) - \varepsilon^2 A(\varepsilon(\cdot - t), \varepsilon^3 t) \tilde{f}_1(0)(\cdot)\|_{H^{2+s}} \leq C\varepsilon^{5/2}.$$

Theorem 5.2 and Theorem 5.3 can be reformulated in a similar way.

#### APPENDIX C. BLOCH TRANSFORM ON THE REAL LINE

In this section we recall basic properties of Bloch transform. Our presentation follows [16]. Bloch transform  $\mathcal{T}$  generalizes Fourier transform  $\mathcal{F}$  from spatially homogeneous problems to spatially periodic problems. Bloch transform is (formally) defined by

$$(28) \quad \tilde{u}(\ell, x) = (\mathcal{T}u)(\ell, x) = \sum_{j \in \mathbb{Z}} e^{ijx} \hat{u}(\ell + j),$$

where  $\hat{u}(\xi) = (\mathcal{F}u)(\xi)$ ,  $\xi \in \mathbb{R}$  is the Fourier transform of  $u$ . The inverse of Bloch transform is given by

$$(29) \quad u(x) = (\mathcal{T}^{-1}\tilde{u})(x) = \int_{-1/2}^{1/2} e^{i\ell x} \tilde{u}(\ell, x) d\ell.$$

By construction,  $\tilde{u}(\ell, x)$  is extended from  $(\ell, x) \in \mathbb{T}_1 \times \mathbb{T}_{2\pi}$  to  $(\ell, x) \in \mathbb{R} \times \mathbb{R}$  according to the continuation conditions:

$$(30) \quad \tilde{u}(\ell, x) = \tilde{u}(\ell, x + 2\pi) \quad \text{and} \quad \tilde{u}(\ell, x) = \tilde{u}(\ell + 1, x)e^{ix}.$$

The following lemma specifies the well-known property of Bloch transform acting on Sobolev function spaces.

**Lemma C.1.** *Bloch transform  $\mathcal{T}$  is an isomorphism between*

$$H^s(\mathbb{R}) \quad \text{and} \quad L^2(\mathbb{T}_1, H^s(\mathbb{T}_{2\pi})),$$

where  $L^2(\mathbb{T}_1, H^s(\mathbb{T}_{2\pi}))$  is equipped with the norm

$$\|\tilde{u}\|_{L^2(\mathbb{T}_1, H^s(\mathbb{T}_{2\pi}))} = \left( \int_{-1/2}^{1/2} \|\tilde{u}(\ell, \cdot)\|_{H^s(\mathbb{T}_{2\pi})}^2 d\ell \right)^{1/2}.$$

Multiplication of two functions  $u(x)$  and  $v(x)$  in  $x$ -space corresponds some convolution in Bloch space:

$$(31) \quad (\tilde{u} \star \tilde{v})(\ell, x) = \int_{-1/2}^{1/2} \tilde{u}(\ell - m, x) \tilde{v}(m, x) dm,$$

where the continuation conditions (30) have to be used for  $|\ell - m| > 1/2$ . If  $\chi : \mathbb{R} \rightarrow \mathbb{R}$  is  $2\pi$  periodic, then

$$(32) \quad \mathcal{T}(\chi u)(\ell, x) = \chi(x)(\mathcal{T}u)(\ell, x).$$

The relations (31) and (32) are well-known and can be proved from the definition (28).

## REFERENCES

- [1] J. L. Bona, T. Colin and D. Lannes. Long wave approximations for water waves. *Arch. Ration. Mech. Anal.*, **178** (2005), 373–410.
- [2] K. Busch, G. Schneider, L. Tkeshelashvili and H. Uecker. Justification of the Nonlinear Schrödinger equation in spatially periodic media. *Z. Angew. Math. Phys.*, **57** (2006), 1–35.
- [3] C. Cercignani and D. H. Sattinger. *Scaling limits and models in physical processes*. DMV Seminar, 28. Birkhäuser Verlag, Basel, 1998.
- [4] M. Chirilus-Bruckner, C. Chong, O. Prill and G. Schneider. Rigorous description of macroscopic wave packets in infinite periodic chains of coupled oscillators by modulation equations. *Discrete Contin. Dyn. Syst. Ser. S* **5** (2012), 879–901.
- [5] M. Chirilus-Bruckner, W.-P. Düll and G. Schneider. Validity of the KdV equation for the modulation of periodic traveling waves in the NLS equation. *J. Math. Anal. Appl.* **414** (2014), 166–175.
- [6] D. Chiron and F. Rousset. The KdV/KP-I limit of the nonlinear Schrödinger equation. *SIAM J. Math. Anal.* **42** (2010), 64–96.
- [7] C. Chong and G. Schneider. The validity of the KdV approximation in case of resonances arising from periodic media. *J. Math. Anal. Appl.* **383** (2011), 330–336.
- [8] F. Chazel. Influence of bottom topography on long water waves. *M2AN Math. Model. Numer. Anal.* **41** (2007), 771–799.
- [9] F. Chazel. On the Korteweg-de Vries approximation for uneven bottoms. *Eur. J. Mech. B Fluids* **28** (2009), 234–252.
- [10] W. Craig. An existence theory for water waves and the Boussinesq and Korteweg-de Vries scaling limits. *Comm. Partial Differential Equations* **10** (1985), 787–1003.
- [11] T. Dohnal, A. Lamacz and B. Schweizer. Bloch-wave homogenization on large time scales and dispersive effective wave equations. *Multiscale Model. Simul.* **12** (2014), 488–513.
- [12] T. Dohnal, A. Lamacz and B. Schweizer. Dispersive homogenized models and coefficient formulas for waves in general periodic media. *Asymptotic Anal.* **93** (2015), 21–49.
- [13] W.-P. Düll and G. Schneider. Validity of Whitham’s equations for the modulation of wave-trains in the NLS equation. *J. Nonlinear Science* **19** (2009), 453–466.

- [14] W.-P. Düll. Validity of the Korteweg-de Vries approximation for the two-dimensional water wave problem in the arc length formulation. *Comm. Pure Appl. Math.* **65** (2012), 381–429.
- [15] W.-P. Düll, K. Sanei Kashani and G. Schneider. The validity of Whitham’s approximation for a Klein-Gordon-Boussinesq model. *SIAM J. Math. Anal.* **48** (2016), 4311–4334.
- [16] T. Gallay, G. Schneider and H. Uecker. Stable transport of information near essentially unstable localized structures. *Discrete Contin. Dyn. Syst. Ser. B* **4** (2004), 349–390.
- [17] T. Iguchi. A long wave approximation for capillary-gravity waves and an effect of the bottom. *Comm. Partial Differential Equations* **32** (2007), 37–85.
- [18] T. Iguchi. A shallow water approximation for water waves. *J. Math. Kyoto Univ.* **49** (2009), 13–55.
- [19] T. Kano and T. Nishida. A mathematical justification for Korteweg-de Vries equation and Boussinesq equation of water surface waves. *Osaka J. Math.* **23** (1986), 389–413.
- [20] L.V. Ovsiannikov. Cauchy problem in a scale of Banach spaces and its application to the shallow water theory justification. Appl. Methods Funct. Anal. Probl. Mech., IUTAM/IMU-Symp. Marseille 1975, *Lect. Notes Math.* **503** (1976), 426–437.
- [21] G. Schneider. Validity and limitation of the Newell-Whitehead equation. *Math. Nachr.* **176** (1995), 249–263.
- [22] G. Schneider. Limits for the Korteweg-de Vries-approximation. *Z. Angew. Math. Mech.* **76** (1996), 341–344.
- [23] G. Schneider and C. E. Wayne. The long-wave limit for the water wave problem. I. The case of zero surface tension. *Comm. Pure Appl. Math.* **53** (2000), 1475–1535.
- [24] G. Schneider and C.E. Wayne. The rigorous approximation of long-wavelength capillary-gravity waves. *Arch. Ration. Mech. Anal.* **162** (2002), 247–285.
- [25] G. Schneider. The role of the Nonlinear Schrödinger equation in nonlinear optics. In Oberwolfach seminars 42: *Photonic Crystals: Mathematical Analysis and Numerical Approximation* by W. Dörfler, A. Lechleiter, M. Plum, G. Schneider, and C. Wiersing. Birkhäuser 2011.
- [26] G. B. Whitham. *Linear and nonlinear waves*. Reprint of the 1974 original. Pure and Applied Mathematics. John Wiley & Sons, Inc., New York, 1999.